DOCUMENT RESUME

ED 039 923 PS 002 924

AUTHOR	Medley, Donald M.
TITLE	OScAR Goes to Nursery School: A New Technique for Recording Pupil Behavior.
INSTITUTION	Educational Testing Service, Princeton, N.J.
PUB DATE	May 69
NOTE	17p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$0.95
DESCRIPTORS	Behavior Rating Scales, Classroom Environment, *Classroom Observation Techniques, *Measurement Instruments, Student Behavior
IDENTIFIERS	Personal Record Of School Experience

ABSTRACT

A new instrument for classroom observation called the Personal Record of School Experience (PROSE), has been developed at Educational Testing Service. While PROSE was invented for a specific project involving 2000 4-year-old ghetto children, its general applicability has already become apparent. PROSE represents a primitive method of classroom observation since it does not provide an observer rating type of classroom record. PROSE strives to eliminate observer subjectivity by requiring only the objective recording of observable events. The record is also personal since the observer only watches one child at a time and records all his activity. PROSE is thus similar to ecological studies, except that the child's record is coded in the classroom rather than from a later review of films and recordings taken in class. The in-class coding is made possible by the PROSE "language," based on 11 word statements that can be fed directly to the computer. Static conditions, such as subject matter, class organization, and instructional materials are also recorded. The 11 word statements used by the observer code all aspects of the particular students's current activity, including level of attention, physical activity, and manifest affect. (MH)

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

RM-69-15

RESEARCH MEMORANDUM

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED039923

OScAR GOES TO NURSERY SCHOOL:

A NEW TECHNIQUE FOR RECORDING PUPIL BEHAVIOR

Donald M. Medley

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Educational Testing Service TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

PS002924

Paper presented at the meetings of the American Educational Research Association, Los Angeles, California, February 8, 1969.

Educational Testing Service
Princeton, New Jersey
May 1969

OScAR GOES TO NURSERY SCHOOL:

A NEW TECHNIQUE FOR RECORDING PUPIL BEHAVIOR

In the time made available to me today I would like to share with you some notions about observational methodology in the context of a new instrument we have developed at Educational Testing Service. This instrument was developed for a particular project, but it has already become clear that it is likely to be useful in other studies and for other purposes.

I have never been able to discuss methodological issues comfortably in the abstract--in fact, I find it almost impossible to think constructively about them unless they arise in connection with a specific research problem. So it is that, when I was invited to participate in this symposium and share with you some of the things we have learned about classroom observation, I found myself turning automatically to our most recent attempt to apply the "OScAR Technique" to a practical research problem. And so it is that today I will be describing some of the constraints and demands placed on us by the nature of the study and some of the devices--the methodological straws, you might say--that we clutched at in meeting them.

Ever since I first became aware of their existence, my own concept of the role of structured observations in research has steadily evolved toward the more primitive. Originally I shared a notion which still has wide currency--the notion that the function of the classroom observer is to generate an evaluative judgment or opinion, usually referred to as a <u>rating</u>, of the teacher of the class. There are, I believe, certain purposes to be served by such ratings. Whenever the opinion of the classroom visitor is

the desired end-product, a rating based on such a visit provides that opinion in a convenient form. But when a measure of the characteristic being rated is needed, ratings are untrustworthy indeed. Evidence that ratings by different observers correlate highly does not necessarily mean they are reacting to the trait they are supposed to be rating. What is needed is evidence that they correlate with some other measure. Attempts to do this with teacher ratings have not been very successful (Medley & Mitzel, 1963, pp. 257-258).

For a number of years many of us have been working under the assumption that if the purpose of classroom observations is to yield measurements of classroom behavior, rather than observers' opinions about teachers or pupils, something different from ratings must be used. If by measurements of behavior we mean numbers which reliably discriminate among the classrooms observed according to the types and amounts of events that occur in them, our observers must record the events themselves rather than opinions derived from them. If classrooms differ significantly with respect to frequencies of behaviors in a certain category, then that category is said to define a behavior dimension and the observational technique is said to measure that dimension.

Implied in this view is the assumption that the constructor of the technique by the act of defining the items or categories that determine which behaviors are to be recorded makes an a priori definition of the dimensions of behavior to be measured by the technique. The verbal category systems devised by Withall (1949) and Flanders (Amidon & Flanders, 1963), for example, were both based on definite hypotheses about what dimensions of teacher behavior are important, and the fact that teachers were found to

differ with respect to the categories supported the theories behind the hypotheses.

In my own work in collaboration with Harold Mitzel (Medley, 1963), we sought to move to a level I have already referred to as more primitive than this. The items or categories we designed for our observers were not intended to identify important dimensions; nor did we require that each one should discriminate among classes--or be of sufficient psychological interest to constitute a meaningful dimension of behavior if it did. Our only requirement was that different observers be able to agree about the frequencies of events in each category. If the number and variety of categories used was sufficient, it should be possible to combine categories (with or without varying weights) into super-categories which would discriminate teachers along psychologically or pedagogically relevant dimensions.

A somewhat homely analogy may serve to clarify this point. Suppose I am sending children to the zoo with the purpose of finding out what kind of animals are housed there. If I used a rating approach I would teach my observers how to recognize lions, tigers, elephants, ostriches, and then ask them to report back to me which of them they saw. If I used the OScAR technique, I would give them a list of characteristics such as number of legs, fur or feathers, color, etc., and ask them to check which characteristics were possessed by each beast they saw. From each child's list I would then infer which animal that child had seen.

In the second approach my conclusions would be much less dependent for their accuracy on the children's ability to recognize animals; indeed, my results might identify some species I had not expected them to see at all. The accuracy of my conclusions would depend partly on the completeness of

my checklist of things to look for, and partly on the children's ability to count legs, discriminate fur from feathers, recognize colors, etc. If these discriminations were primitive enough, based on obvious cues, and if my list were sufficiently varied and complete, my conclusions should approach 100% accuracy.

Similarly, if I ask a classroom observer only to make discriminations based on simple overt behaviors, and if I can construct a sufficiently comprehensive list of them for him to record, it becomes possible to secure descriptions of dimensions of behavior virtually independent of the training or experience of the observers. Indeed, there is no reason why an observer should even know what dimensions are being measured. This is the sense in which I refer to the OScAR technique as more primitive than category systems like those of Withall and of Flanders already cited.

Before turning to the specific application I wish to discuss, let me emphasize the fact that any individual item on a schedule may contain very little information, but a combination of several occurring together can pinpoint an important piece of data. To know only that an animal has hoofs, or fur, or stripes does not tell us much. To know that he has all three makes us pretty nearly certain that he is a zebra. It is by piling up a number of simple facts that we develop useful information.

With this as background I would like to describe a new observation system we have designed, one in which we have tried to adapt these ideas to a new study.

The study in question is a large-scale longitudinal study of early childhood education. A sample of 2,000 four-year-old children from ghetto areas will be identified, studied intensively for a five-year period

ending when most of them are in the third grade. Particular emphasis is to be placed upon the experiences each individual child has in school, including any nursery school or Head Start center he may attend. One purpose of the study is to identify characteristics of these early classroom experiences which are associated with success in school by the time the child reaches the third grade.

The requirements of the study I have mentioned set some important constraints on us. First, the sheer size of the project is a factor to be considered. The proposal is to secure a minimum of 20 observations per year of each of 2,000 children for five years. That means 40,000 records in the first year, and (allowing for attrition) something like 150,000 total.

Second, we are being asked to observe pupils instead of teachers. Most of our past experience has been gained in observing teachers--a much easier task. For one thing, there is only one teacher in a class, as a rule--and his role is central. In that same classroom there are between 15 and 45 pupils, whose roles vary from central to peripheral, and who do not always act as if they knew their roles.

A third constraint arises from a decision to use as observers people recruited in the community, or at least in the subculture being studied. The middle-class white observer has never been fully accepted in the ghetto--and in these times of social change is even less so. And even when he is, there is serious question whether he may not be so alien to the culture under study that his perceptions are distorted beyond the limits of usefulness. This made it necessary to design an instrument which can be used by

PS 002924

observers who have no background either in psychology or pedagogy, and no understanding of or particular commitment to the scientific study of behavior--perhaps, even with little or no formal education beyond high school.

Fourth, a study like the one under discussion, designed to obtain comprehensive (if not exhaustive) data about a sample, requires something more like a descriptive record than a profile of measurements. Full exploitation of the potential of this study requires as comprehensive and many-faceted a record of the child's experience as can be obtained. It must be possible to look at the data from many different angles, to test many different hypotheses--some of which, although today we do not suspect their existence, may seem of central importance five years from now.

There were a number of other contraints placed on the instrument to be constructed (some of which did not become apparent until they were encountered during the developmental process), but these were the major ones. Taken as a whole, they seemed to call for an instrument which would describe classroom behavior as it related to the individual pupil, comprehensively, objectively, and in a readily quantifiable form--and an instrument that a person completely inexperienced in structured observations could learn to use easily and accurately.

Let me turn now to the instrument we have developed, which we call the Personal Record of School Experience, or PROSE, and try to show how we have attempted to meet these requirements.

The first thing to be noted is that the PROSE record is personal--the PROSE recorder focuses on one pupil at a time and records only what happens

to him. If the pupil attends to the teacher, for example, the teacher behavior at that time is part of the child's experience and is recorded; if the pupil does not attend to the teacher, the teacher's behavior at that time is not part of the child's experience and is not recorded.

If the behavior of each pupil in a class is sampled in this way, the behavior of the whole class has been sampled, and a description of the class as a whole may be constructed by combining individual records. But this process cannot be reversed. It is not possible to observe everything that happens in a class; and even if we could do so, we could not infer that all of it happened to the child we are interested in studying.

As a matter of fact, one soon finds that it is not even practical to observe and record everything that happens to a single child. There is one group that claims to be able to do this--the ecological psychologists. Using the latest technological aids, the ecologist claims to be able to record everything that happens to a child, either by dictating a running narrative account into a tape recorder or by photographing the child's behavior on film or videotape.

Whether an ecological report is complete or not, it is quite useless from a scientific point of view until it has been coded into quantitative form--or, as I prefer to phrase it, until it has been translated into a language a computer can process. The task of coding such records is, for all practical purposes, the same task as the one that the classroom observer performs in the classroom. The coder working with the ecological record has one advantage over the classroom observer--the data he is

coding are recorded in permanent form, so that he can spend all the time he needs on the coding task. He can read and reread a transcript, view and review a film or videotape recording until he has it coded just right. The classroom observer, on the other hand, must work in real time; he must code behaviors that appear once, then vanish forever. There is neither time to ponder a difficult decision nor opportunity to go back and reconsider. "The moving finger writes, and having writ, moves on."

But the coder working from an ecological record must work with incomplete data; the cues the camera or recorder may have missed can never be taken into account. On the other hand, the classroom observer has access to everything that happens (on camera or off), whether the ecological recorder saw it as important enough to record or not. If behaviors are coded by our observers in the classroom, it takes one hour of coder time to code one hour of behavior. If behaviors are coded from ecological records, it takes one hour to record one hour of behavior, and at least one more hour to code it. If any advantage is taken of the capability to rerun tapes or reread verbal records, it may take two or more hours to code one hour of behavior. Thus quantifiable behavior records based on ecological records may use anywhere from two to four or five times as many man-hours per hour of behavior as are needed when behavior is coded "live."

If an equal number of man-hours were invested in observing behavior and coding it "live" rather than in coding and recoding the same records, from two to five times as much data could be obtained with the same investment of time and money! Since our objective is to describe classroom behavior on the basis of observations of relatively small samples, we must balance the larger observer errors made when behaviors are coded live against the

larger sampling errors incurred by ecological coders using a much smaller behavior sample. It is our experience that sampling errors tend to be much larger than the observer errors, so we have opted for in-class coders.

The ecological record did, however, attract us as an ideal approach except for this one limitation--the inefficient use of time. We decided, therefore, to attempt to design a technique by which the classroom observer might compile a record analogous to an ecological one but precoded, that is, an "ecological" record written in a language the computer could read, analyze, and interpret.

Achieving this goal involved the accomplishment of three steps: (1) the development of a vastly simplified "language"--simplified both as to vocabulary and syntax; (2) the development of a way of writing the language so that the computer could read it; and (3) the development of a procedure by which the classroom observer could record behaviors in that language accurately and easily.

The language we developed is based on 11-word statements made up by the observer to describe events observed in the classroom. Each word in a statement is a position in the statement which may be filled by one of two or more alternatives. The 11 words are represented on the front (or statement) side of the recording form (see Figure 1) by numbers one to 11, each followed by the alternative choices for that word. Opposite each choice are 15 mark-sensing spaces, numbered to correspond to events numbered 1 to 5 in three cycles of five events each.

The procedure may perhaps best be understood in terms of an example. Suppose that the event to be recorded is one in which the teacher is showing

the class how to plant a seed in a flowerpot, and that the pupil on whom the observation is focused is watching attentively. The task of the observer is, simply, to mark whichever alternative (if any) to each of the 11 words best describes this event.

On the first word, for example, he marks "1.PART" in Box 1 (because this is event number one) to indicate that the child was a part of a group to which an adult was attending. On the second word he marks "2.TCHR" (again in Box 1) to indicate that the adult in question was the teacher. On the third word he marks "3.SHTL" indicating that the teacher was showing or telling the pupils something.

Words four and five are omitted as not applicable (they are used only for events involving peer interactions) and on the sixth word "6.MTL" is marked, to indicate that the event involved materials of some sort. On word seven the recorder indicates the sex and ethnic group mix in the event. If, for instance, the pupil is of the same sex and ethnic group as the teacher, "7.SSSG" is marked.

The eighth word is used to describe the level of attention the pupil is paying to the adult; in this case, conforming to the role expected of a pupil--i.e., paying attention--so "8.COOP" is marked. Word nine is used to indicate physical activity on the part of the pupil being observed. Since this pupil is not moving, word nine is left blank. The tenth word is used only when the pupil is alone, and the eleventh is used to indicate manifest affect; both are left blank in this example.

Since events in the classroom occur rather rapidly, it is impossible for one observer to record all of the things that happen to a child during

the time he is under observation. Therefore, only a sample of such events is recorded, a sample which is determined by a small timing device which the observer wears. This device emits a signal at fixed intervals (we are currently using 25-second intervals) into a hearing-aid type earphone in the observer's ear.

The observer locates the pupil to be observed, starts the timing device, and watches what happens to the pupil. When the timer sounds, he is to record a statement describing the event occurring at that moment. As soon as the statement has been recorded, the observer begins to watch the pupil again in preparation for the next signal. Five statements describing five events complete one cycle.

In addition to the 11-word statements used to describe transitory characteristics of each event, the language also provides means for describing more stable conditions; this information is recorded on the reverse or context side of the instrument (see Figure 2).

Static conditions include the subject being taught, the way in which the class is organized, where in that organization the pupil fits, the general climate of the class, materials used by the child, and his location in the classrooms. Since all of these things are relatively stable, they are recorded only once for each cycle of five events. After the observer records the fifth statement in a cycle, he turns the form over and marks those alternatives to words 1-5 and 8 which best describe the context in which the five events described on the statement side took place.

Continuing the same example, we mark "1.SCI" to indicate that the teacher is giving a science lesson; "2.SHOW" and "2. TELL" (in Box T) to

indicate that the teacher (T) is demonstrating and explaining; "3.MID," which would indicate that the child is in the middle of the group; "4.ALL" (in Box Ad) to indicate that the group in question contains all of the pupils (and an adult); and "8.ASSIGNED SEAT" (in Box L) to indicate that the pupil is in his regular seat.

If one were to read this record without having seen the event, he would note that the child was quietly sitting in the middle of the class in his own seat while the teacher was demonstrating something about science. While this does not describe exactly what happened, it describes the essential nature of the event in a way analogous to (though less detailed than) an ecological record.

When an observer visits a class he has with him one PROSE form for each pupil in the class, arranged in random order. He completes one cycle of five statements on each pupil in turn, then another set of cycles, and finally a third (if time permits). In the longitudinal study we are planning, 100 observations of each child will be made, comprising 1,000 statements describing samples of his behavior observed on 100 different days throughout the five years of the study.

The number of different statements that can be made is very large. Over 200,000 different sensible statements can be made up using only the 11 words on the statement side. Since each statement can also be modified according to the static conditions under which the event occurs, we might say that the instrument has over a million items on it. Or at least, that it is possible to record much more detail in a PROSE record than could be used in any normal statistical analysis. It is in this wealth of detail that a PROSE record resembles an ecological one. The difference is that with the

aid of SCRIBE, the high-speed test-scoring machine at ETS, all of this information may be read into a computer without further human intervention. Ultimately, as I have mentioned, the computer will have access to a sort of narrative description, 1,000 statements long, of the five-year school experience of each child in the study.

The PROSE records on each child will be accompanied by a vast amount of other information about him--environmental, developmental, physical, social--to which his experiences in school may be related. As a hypothesis is developed about such a relationship, it should be possible to find in the pupil's PROSE records whatever information about his school experiences is relevant to it. It would then be a straightforward problem in computer programming to extract that information in quantitative form and perform whatever statistical analyses are needed. But that is another problem to be dealt with another day.

References

Amidon, E. J., & Flanders, N. A. The role of the teacher in the classroom. Minneapolis, Minnesota: Paul S. Amidon and Associates, Incorporated, 1963.

Medley, D. M. Experiences with the OScAR technique. Journal of Teacher Education, 1963, 14, 267-273.

Medley, D. M., & Mitzel, H. E. Measuring classroom behavior by systematic observations. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago, Illinois: Rand McNally, 1963. Pp. 257-328.

Withall, J. Development of a technique for the measurement of socio-emotional climate in classrooms. Journal of Experimental Education, 1949, 17, 347-361.

Be sure each mark is *dark* and *completely fills* the answer space.
Do not make any stray marks on either side of this sheet.

PERSONAL RECORD OF SCHOOL EXPERIENCE

A B C D E F G H I J K L M N O

1. INIT STAR PART LSWT
2. TCHR AA TAA OBS OTH
3. POS PRM SHTL LSQU
DO4 CNTR NEG
4. AGR INIT COOP WTHD RST
5. AGR INIT COOP WTHD RST

CYCLE A CYCLE B CYCLE C

6. CNTC MTL VRB
7. OSOG OSSG SSOG SSSG
8. COOP DSTR RIS WOA DSRP
9. HIWL MDWL HINL MDNL LOW
10. FANT DVG CVG WRK KIN
11. POS NEG

REMARKS

DO NOT WRITE BELOW THIS LINE

Figure 1. Statement Side of PROSE Form

Educational Testing Service
Princeton, N. J. 08540

B108P50D

PERSONAL RECORD OF SCHOOL EXPERIENCE

1. ARITH	A	B	C
ARTCRF	A	B	C
EXGM	A	B	C
FRPL	A	B	C
HLTHYG	A	B	C
LANG	A	B	C
MUSRTH	A	B	C
RSTSNK	A	B	C
RTNTRN	A	B	C
SCI	A	B	C
SENS	A	B	C
SOSK	A	B	C
SOST	A	B	C
3. NEXT	A	B	C
NEAR	A	B	C
MID	A	B	C
FRNG	A	B	C
OUT	A	B	C
5. ATT EXC	A	B	C
ATT TNS	A	B	C
NSY EXC	A	B	C
NSY BSY	A	B	C
QU BSY	A	B	C
QU IDL	A	B	C

	A		B		C	
2. SHOW	T	O	T	O	T	O
TELL	T	O	T	O	T	O
LEAD	T	O	T	O	T	O
DSCS	T	O	T	O	T	O
PEER	T	O	T	O	T	O
MNG	T	O	T	O	T	O
SPVS	T	O	T	O	T	O
RSRCE	T	O	T	O	T	O
HSKP	T	O	T	O	T	O
INDATT	T	O	T	O	T	O
LSWT	T	O	T	O	T	O
NCNT	T	O	T	O	T	O

	A		B		C	
4. ONE	Ad	NA	Ad	NA	Ad	NA
TWO	Ad	NA	Ad	NA	Ad	NA
3-5	Ad	NA	Ad	NA	Ad	NA
6+	Ad	NA	Ad	NA	Ad	NA
ALL	Ad	NA	Ad	NA	Ad	NA

6. USED NUMBERS	A	B	C
USED WORDS	A	B	C
SNG, TKD TO SLF	A	B	C
HELPED OTHER P	A	B	C
COMFORTED P	A	B	C
SHAFF P	A	B	C
ASKED P FOR HELP	A	B	C
REC'D HLP, AFF	A	B	C
REJECTED BY GRP	A	B	C
ACCIDENT, HURT	A	B	C
SHOWED FEAR	A	B	C
CRIED	A	B	C
LOST TEMPER	A	B	C
TATTLED	A	B	C
LED OTHER P	A	B	C
BOSSED OTHER P	A	B	C
WRECKED SOMETHING	A	B	C
REFUSED HELP	A	B	C
RESISTED AD	A	B	C
DISOBEYED	A	B	C
SHOST TO AD	A	B	C
RID	A	B	C
BAD EX	A	B	C
GOOD EX	A	B	C
CHORE, ERRAND	A	B	C
SHAFF AD	A	B	C
WAITED	A	B	C

	A		B		C	
7. CLD4QT	P	W	P	W	P	W
THRT	P	W	P	W	P	W
LSTMP	P	W	P	W	P	W
LFTW	P	W	P	W	P	W
PHRSTR	P	W	P	W	P	W
8. ART (PAINT, CLAY)	M	L	M	L	M	L
CRAFT	M	L	M	L	M	L
MUSICAL INSTR	M	L	M	L	M	L
BOOKS, WRITING	M	L	M	L	M	L
PUZZLES, QU GAMES	M	L	M	L	M	L
BLOCKS	M	L	M	L	M	L
WHEEL TOYS	M	L	M	L	M	L
TOOLS, WK BENCH	M	L	M	L	M	L
SAND	M	L	M	L	M	L
GYM, EXERCISE	M	L	M	L	M	L
PET	M	L	M	L	M	L
CLOTHES, JEWELRY	M	L	M	L	M	L
FOOD, WATER	M	L	M	L	M	L
DOLLS	M	L	M	L	M	L
HOMEMAKING	M	L	M	L	M	L
CLEANUP TOOLS	M	L	M	L	M	L
AUDIO-VISUAL DEV	M	L	M	L	M	L
SP INSTRC DEV	M	L	M	L	M	L
SCIENCE EQUIP	M	L	M	L	M	L
ASSIGNED SEAT		L		L		L
TEACHER'S DESK		L		L		L
OPEN AREA		L		L		L
TOILET		L		L		L
PLAYGROUND		L		L		L
OUT OF ROOM		L		L		L

Figure 2. Context Side of PROSE Form